Limits of Effective Altruism

Jonathan Berman, London Business School, UK
Alixandra Barasch, University of Pennsylvania, USA
Emma Levine, University of Pennsylvania, USA
Deborah Small, University of Pennsylvania, USA

Charitable giving could do the most social good if people allocated donations to the most effective charities. We find that consumers believe that giving to charity is a subjective decision, and as a result, they prioritize their personal preferences at the expense of maximizing effectiveness.

[to cite]:

[url]:
http://www.acrwebsite.org/volumes/1019461/volumes/v43/NA-43

[copyright notice]:
This work is copyrighted by The Association for Consumer Research. For permission to copy or use this work in whole or in part, please contact the Copyright Clearance Center at http://www.copyright.com/.
The Science of Charitable Giving and Pro-Social Behavior

Chair: Indranil Goswami, University of Chicago, USA

Paper #1: On Being the “Tipping Point”: Threshold Incentives Motivate Behavior
Lalín Anik, Duke University, USA
Michael Norton, Harvard Business School, USA

Paper #2: “I feel your pain” The efficacy of instantiating states in charitable appeals
Carey Morewedge, Boston University, USA
David Carlin, Williams College, USA

Paper #3: The Effect of Default Amounts on Charitable Donations. Evidence from a Large-Scale Field Experiment
Indranil Goswami, University of Chicago, USA
Oleg Urminsky, University of Chicago, USA

Paper #4: Limits of Effective Altruism
Jonathan Berman, London Business School, UK
Alixandra Barasch, The Wharton School, USA
Emma Levine, The Wharton School, USA
Deborah Small, The Wharton School, USA

SESSION OVERVIEW
Recent data from the National Center for Charitable Statistics shows that almost 3/4 of the contributions to nonprofit organizations come from private individuals. However, even with such high stakes, the charitable solicitations and other pro-social appeals used to solicit individual donations are rarely determined by empirical evidence. Therefore, it is important to systematically and scientifically investigate the various levers that can be used to influence pro-social behavior. In this session we bring together research both from the lab and from the field to shed light on the psychology of donation decisions and increase the efficacy of charitable appeals. The research in this session investigates the interplay between the structure of the appeal, information provided, and the social context of giving. We believe that these findings not only offer significant practical implications, but also leverage the setting of charitable giving to make important theoretical contributions for decision researchers.

Solicitation campaigns aiming to increase participation often employ some kind of incentive. Anik and Norton propose that highlighting the crucial role that an individual decision maker can play to tip a contingent reward to a higher level, affecting not only themselves but also other donors, can increase their motivation to participate. For example, informing a potential participant that getting a blood test would increase the amount donated for every participant from a baseline of $x to a higher level of $y, increased the likelihood of participation. The effectiveness of tipping points is driven by a sense of responsibility towards fellow participants. The intervention performs better than offering non-contingent financial rewards and is more effective at lower levels of contingent reward than at higher levels. This suggests that the intervention could be not only behaviorally effective but also economically efficient.

Morewedge and Carlin propose another way to increase participation by instantiating the psychological state of the charity’s beneficiaries. This research demonstrates that empathy-invoking solicitations that invoke negative or unpleasant emotions to help the donor “feel the pain” of the beneficiary can be more effective than positive appeals, even when subjects evaluate both the appeals as equally important. Appeals that increase potential donors’ empathy for the targets can also potentially explain previously documented phenomena like the identifiable victim effect (Small, Loewenstein, & Slovic, 2007) and martyrdom effect (Olivola & Shafir, 2013) and therefore has both theoretical and practical implications.

Extant research has documented that pre-set defaults have robust effects on choices, in consequential domains, including organ donation and retirement savings. However, the effect of an active-choice default on amount (as in charitable solicitations), which involves both the participation rate and average contribution, has not been systematically investigated. Using a large-scale natural field experiment, Goswami and Urminsky find that the effect of defaults on donation decisions is driven by two separate effects—designating a low amount as default lowers the bar and motivates more participants to contribute. However, participating donors scale back their contribution amount when faced with lower defaults. The net effect of these two conflicting forces determines the effect of defaults on donation amounts. The predictions of existing theories of defaults, suggesting either a direct default effect or a backlash against high defaults are both unsupported, suggesting the need to revisit our understanding of such useful nudges, particularly in the context of charitable giving.

Finally, Berman, Barasch, Levine, and Small examine how consumers make decisions regarding charities when effectiveness information (i.e., the total of good done per dollar donate) is available. The authors find that because consumers tend to believe that the choice of where to donate money is a subjective decision, they often weigh personal preferences at the expense of maximizing effectiveness. Specifically, they find that that consumers are less likely to use effectiveness information for charitable decisions than for investment decisions, and further, consumers are less likely to search by effectiveness ratings as compared to investments. They also find consumers use less effectiveness information when choosing between charities in a diverse rather than uniform choice set.

Taken together, we expect that this session will be of interest to a broad audience, facilitating connections among researchers interested in either pro-social behavior, judgment and decision making, persuasion and social influence, incentives or policy “nudges”. This session will also contribute to the diversity of research methods, featuring data from varied methodologies. We believe that understanding how decision makers think when confronted with a charitable solicitation identifying the conditions that yield optimal effectiveness is of crucial importance for both theoretical and practical reasons. We hope that this session will advance these connections and promote future research that improves our understanding of the science of charitable giving.

On Being the “Tipping Point”: Threshold Incentives Motivate Behavior

EXTENDED ABSTRACT

We explore and document a new intervention to motivate behavior: being the tipping point, the person whose contribution passes a threshold that kicks in a higher reward for the self and others. Anik, Norton, and Ariely (2014) showed that offering incentives contingent on other people’s behaviors is motivating: “If 75% of people engage in this behavior, we will increase the rewards for everyone.” Across three studies, we explore whether specifically being the person who tips that incentive is particularly motivating: “We are currently at 74% and you will bring us up to 75%.”

Advances in Consumer Research
Volume 43, ©2015
In Study 1, we asked participants to imagine a hypothetical scenario where donation would be made to a local health charity for each person who got a blood test. Participants were randomly assigned to a 2 (condition: tipping vs. control) × 3 (level of progress: low, medium, high) design. We told those in the tipping conditions that [9, 49, 89] people have gotten a blood test so far, and the amount to be donated per participant was [25¢, $1.27, $6.41]. If they agreed to get a blood test, since they would be the [10th, 50th, 90th] person, each donation would increase to [37.5¢, $1.90, $9.61], respectively. If they decided not to get a blood test, the per-participant donation rate would stay at [25¢, $1.27, $6.41]. We told those in the control conditions that [10, 50, 90] people have gotten blood tests so far, and the amount donated per person was [37.5¢, $1.90, $9.61]. If they also got a blood test, they would be the [11th, 51st, 91st] person, and each donation would still be [37.5¢, $1.90, $9.61]. Next, participants indicated their likelihood of getting a blood test using a 7-point scale (1 = not at all likely, 7 = definitely likely).

A 2 (condition: tipping, control) × 3 (level of progress: low, medium, high) ANOVA revealed a significant main effect of condition, \( F(1, 286) = 17.41, \ p < .001 \). The main effect for level of progress and the condition X progress interaction were both marginally significant, all Fs > 2.52, ps = .08. Participants in the tipping conditions indicated a significantly higher likelihood of getting a blood test (\( M = 5.37, SD = 1.81 \)) than those in the control conditions (\( M = 4.44, SD = 1.91 \)), \( t(290) = 4.30, p < .001 \), suggesting that overall being the tipping point was indeed more motivating. A closer analysis revealed that at low and medium progress levels, participants’ likelihoods of getting a blood test were significantly higher when they were the tipping point, all ts > 3.16, all ps < .003. However, for high progress level, being the tipping point was not more motivating than the baseline, \( t(108) = .64, p = .53 \).

One reason tipping points are more motivating when they are earlier, rather than later, could be that large contingent monetary incentives might have “crowded out” motivation (Deci 1975; Lepper, Greene, & Nisbett 1973). Therefore, in Study 2, we pit the impact of being the tipping point against small, non-contingent monetary incentives. The scenario used was very similar to the one in Study 1. Participants were randomly assigned to one of five conditions. The tipping condition informed participants that they could be the 10th person who tipped the donation from 50¢ to $1 for everyone. Participants in the non-tipping condition were told that they could be the 10th person, and the donation would still be 50¢ per head. Finally, those in one of three non-contingent incentives conditions were not told about other’s behavior, and were informed that $1, $5.50 or $10 would be donated to a charity on their behalf if they got a blood test. Participants indicated their likelihood of testing (1 = no, 0 = yes).

We replicated the effect that social incentives were more effective than non-social incentives (\( M = 75.47\% \) vs. \( 58.67\% \)), \( \chi^2(1) = 3.98, p < .05 \). Further contrasts revealed that percent of participants who agreed to get a blood test in the tipping condition (\( M = 85.71\% \)) differed from the non-tipping condition (\( M = 74.57\% \)), \( \chi^2(1) = 3.36, p < .07 \), as well as from all three non-contingent conditions, \( \chi^2(2) > 4.49 , ps < .04 \). Results show that being the tipping point in a social group is motivating above and beyond offering monetary incentives alone or providing only social information.

In Study 3, we ran a study with real consumer behavior where participants reviewed a restaurant experience in exchange for donations to a project that provided lunchboxes to kids. There were three conditions: a non-contingent incentives condition offered a $3 donation; a tipping condition grouped participants with two others, and by reviewing the restaurant, donations for all three participants would jump from 50¢ each to $1 each (a total of $1 donated in total if the participant didn’t respond to $3 total if they did); and finally, a social no-tipping condition, where the amount donated for each participant was $1, $3 in total for the charity. Using a 7-point scale (1 = not at all, 7 = a great deal), participants answered a series of questions about how much responsibility they experienced towards other participants and towards the kids who would receive the lunchboxes and their feelings of altruism, impact, substitutability and guilt in general. Our dependent variable was the percent of participants who provided a review.

Again, we found that setting incentives contingent on social behavior led to a significantly higher percent of reviews than offering non-social incentives, (\( M = 42.73\% \) vs. \( 33.83\% \)), \( \chi^2(1) = 2.13, p < .04 \). Further analysis showed that the effect of social incentives was driven by the tipping condition as it resulted in participation significantly higher than in social no-tipping condition (\( M = 50.00\% \) vs. \( 35.71\% \)), \( \chi^2(1) = 2.15, p < .04 \). Finally, a mediation analysis revealed that the effect of being the tipping point on participation was mediated only by participants’ feelings of responsibility towards fellow participants, rather than feelings toward the recipient or general feelings of altruism, guilt, substitutability or impact.

“**I Feel Your Pain**” The Efficacy of Instantiating States in Charitable Appeals

**EXTENDED ABSTRACT**

We examine the efficacy of charitable appeals that instantiate the psychological state of the target (the person helped by the charitable organization) in prospective donors in four experiments. Participants were more likely to donate a bonus to the same charity when presented with a more negative but empathy-evoking solicitation than a more positive informational solicitation stressing the importance of its work (Study 1). Participants were more likely to report they would participate in a fundraising event that was semantically related to the experience of the target than a semantically unrelated event (Study 2). Participants were more likely to report they would participate in fundraising events for causes of equal importance when the salient target of causes were possible to empathize with (people) than impossible to empathize with (abstract causes; Study 3). Finally, participants were more likely to report they would participate in a fundraising event for the same target when the event instantiated the state experienced by the target than when the event evoked an unempathetic state (Study 4). The results suggest a mechanism to increase charitable giving and that may help to explain the identifiable victim (e.g., Small, Loewenstein, & Slovic, 2007) and martyrdom effects (Olivola & Shafir 2013) that have been previously identified.

In Experiment 1, we assessed the effect of empathic (evocative) appeals on donation size. In this experiment, one cause (the Make-A-Wish Foundation) was selected and participants were assigned to an empathetic or an informative appeal. Participants in the empathetic appeal condition were given a salient target, an involved description of the suffering of the target, and the role of the charity in helping the target. In the informational condition, participants were provided with relevant information about the charity, their cause, who they help, what some of their events have included, and its fiscal responsibility and outside ratings. Both conditions were similar with respect to the amount of material they provided; the empathetic appeal was more negative emotion inducing and positive emotion reducing in a pretest, \( ts(100) > 3.29, ps \leq .01 \). After reading an appeal, participants received a $1 bonus and donated what they wished of it to that charity. Participants in the empathetic appeal condition donated sig-
nificantly more ($M = 32.2\,\text{¢}; SD = 29.4$) than did participants in the informational appeal condition ($M = 18.3\,\text{¢}; SD = 28.8$) and reported significantly higher levels of empathy for the target ($M_{\text{empathic}} = 6.04$, $SD = 1.50$; $M_{\text{informational}} = 5.25$, $SD = 1.87$), $t(100) = 2.05$, $p = .04$, and $t(100) = 2.36$, $p = .02$, respectively. Neither group deemed the cause more important, $t(100) = 1.32$, $p = .19$.

In Experiment 2, we compared the efficacy of participation in charitable events that would evoke empathy in the donor for targets (semantically related) to events that were pleasurable but unlikely to evoke empathy in the donor for targets (semantically unrelated). More specifically, in Experiment 2, we selected four causes (e.g., quality of life for nursing home residents) and then chose events that had been rated as semantically relevant or irrelevant to those causes (e.g., a bingo night vs. Texas hold'em night). Participants first rated their likelihood of attending all eight events (if they were not for charity). For the relevant condition, for each of the four causes, participants were shown a semantically related event to the cause and rated their likelihood of attending to support that cause. For the irrelevant condition, for each of the four causes, participants were shown a semantically unrelated event to the cause and rated their likelihood of attending to support that cause. Participants in the relevant condition reported feeling more empathy and connection to the targets, $t(116) > 4.03$, $p < .001$, and exhibited a greater increase in willingness to attend the events ($M_{\text{attendance}} = 1.14$, $SD = 1.50$) than participants in the irrelevant condition ($M_{\text{attendance}} = .58$, $SD = 1.40$), $F(1, 116) = 4.24$, $p = .04$.

In Experiment 3, we further tested the proposed process by presenting participants with a charitable fundraising event that was relevant to targets, but manipulating whether or not the salient target of the charity was capable of evoking empathy. Participants first rated their likelihood of attending six events. For a set of three causes, each participant then rated their likelihood of attending three of those events in support of related abstract targets (e.g., clean rivers) or related human targets (e.g., education for underserved school children), so that participants would be more likely to empathize with the target in the latter than in the former condition (Jenni and Loewenstein 1997). We found that participants exhibited heightened empathy and were more willing to participate in the event when the targets were human ($M_{\text{attendance}} = 1.04$, $SD = 1.25$) than abstract ($M_{\text{attendance}} = .34$, $SD = 1.02$), $t(138) = 3.63$, $p < .001$. In contrast, they rated both kinds of causes as equally important and cause-event pairings to be equally relevant across the two conditions, $t(138) = 1.26$, $p > .21$.

In Experiment 4, we manipulated the degree to which events elicited empathic understanding for targets. Participants were presented with semantically related charitable fundraising events for the same three targets (e.g., the hungry), which were designed to either evoke the state experienced by the target in the prospective donor (e.g., giving up lunch for a day) or designed to evoke a state inversely related to the state experienced by the target (e.g., a tasting menu at a Michelin star restaurant). In support of our hypothesis that understanding the target’s state and “feeling his/her pain” drives empathy and charitable commitment, we found that the empathy-evoking events did evoke more empathy for the target, $F(1, 142) = 26.09$, $p < .001$, and increased the likelihood of participation in the fundraising event ($M_{\text{attendance}} = .76$, $SD = 1.07$) relative to the non-empathetic event condition ($M_{\text{attendance}} = -.28$, $SD = .93$), $F(1, 142) = 37.47$, $p < .001$. Events considered, however, did not influence the perceived importance of the causes, $F < 1$.

The Effect of Default Amounts on Charitable Donations. Evidence from a Large-Scale Field Experiment

EXTENDED ABSTRACT

When factors in the decision environment influence the choices people make, changes in those factors (“nudges”) can be used strategically to influence people’s choices. However, the research underlying nudges tends to point to the existence of psychological phenomena, rather than quantifying the effects of such nudges or identifying the optimal configuration. In this paper we investigate how an active-choice default option, a widely studied and much discussed “nudge”, affects donations when used in fundraising appeals, and the implications for optimal choices of default amounts. Using a large-scale field experiment, we find evidence for two primary influences of defaults on donations: a “scale back” effect (lower donation amounts when a small donation amount is defaulted), and a “lowered bar” effect (more people donating when the small amount is defaulted). The interplay of these two effects determines the net effect of defaults on the amount raised from campaigns using a defaulted amount.

We conducted an experiment in the spring 2014 phase of the annual alumni fundraising campaign of a large Midwestern business school. In total, up to two mailings were sent to the 7917 prior donors who had not yet donated that year. Subjects were randomly assigned to one of the cells of a 4 (default level: Low, Medium, High, None) x 2 (reminder about prior donation amount: No, Yes) x 2 (number of options to designate what their donation would be used for: 2, 5) full-factorial design. The default amount was operationalized by highlighting a particular menu option with a different color and labelling it as “suggested”. The menu options presented to the participants were customized based on prior donation behavior, with the middle menu-option being their previous donation amount. In total, the mailings yielded 76 donations, a 0.96% participation rate. The relatively low donation rate was due to targeting “colder” prior donors, who had not responded to appeals in the previous waves of that fiscal year’s campaign.

Setting the low option as the default increased revenue compared to the no-default control ($M_{\text{Low Default}} = $ 2.60 vs $M_{\text{Control}} = $1.87), but that was not the case for the other default amounts. The medium and high defaults yielded similar revenue ($1.75 and $2.11, respectively) as control. Importantly, setting the high menu option as the default did not result in a revenue-reducing backlash effect in this context. These findings were confirmed in a series of regression analyses predicting the log of per-person revenue, controlling for lasso-selected covariates to avoid over-fitting. The effect of defaults on net revenue was robust across demographic and prior donation behavior differences, and was not moderated by the two other experimental manipulations (reminder of prior donation and the number of allocation options).

The effect of defaults on net revenues can be decomposed into the effect of defaults on participation and on average contribution among those donating. The low default significantly increased donation rates, relative to no default (1.6% vs. 0.7%, $c^2 = 10.23$, $p = .001$), and compared to medium and high defaults combined, low defaults resulted in a marginal increase in participation (1.6% vs. 1%, $c^2 = 3.06$, $p = .08$). Therefore, lowering the bar, by setting a lower default, encouraged more donors to participate in the campaign. The lower default may allow the donor to pay less without harming their pro-social self-image (Gneezy et al. 2010) and still experience the same good feeling as donating a larger amount, in effect obtaining “warm glow” (Andreoni 1990) at a discount.
Unlike participation, our inferences about average donations are under-powered, on account of the low number of donation incidences observed. Overall, the average amount donated was lower when an option was defaulted (M = $178), compared to the no-default control (M = $283). Among, the default conditions, the average donations were consistent with a “scale back” effect: highest for the high default (M = $204), $25 lower for the medium default (M = $181), and approximately $20 lower for the low default (M = $162). Therefore, increasing the default amount resulted in participants scaling back their donation amounts. This effect might arise from inferential reasoning about the norms or needs conveyed by the pre-selected default level. Additional lab-based studies demonstrate the “scale back” effect with actual choices.

Our findings did not provide support for the most optimistic prediction, a direct-default effect, in which simply setting one option as a default would consistently increase funds raised. However, we also found no support for the most pessimistic prediction, a backlash effect, in which setting a default (particularly a high amount default) would lead to rejection of the defaulted amount and lower revenue. The conflict between increased participation and lower donation amounts when using a low default points to both the difficulty of blanket policy prescriptions and the limitations of existing theories. In sum, the net effect of defaults on revenue is determined by the relative strengths of the “scale back” and the “lowered bar” effect. We are continuing to track donations, and will be able to also report the long-term effects of defaults on the subsequent year’s donations at the time of the conference.

Limits of Effective Altruism

EXTENDED ABSTRACT

Recent years have seen a rise in the “Effective Altruism” movement, which attempts to quantify the effectiveness of charitable organizations in order to better allocate resources to those in need. The Effective Altruism movement has gained followers from fields like economics and finance, which emphasize maximizing returns. Effective Altruism treats charitable giving as a social investment where the goal is to maximize social return, by donating to causes where one can do the most social good for each dollar donated (Baron and Szymanska 2011; Singer 2009).

We investigate how everyday consumers make decisions about where to donate their money. We argue that consumers do not naturally think about donation decisions as similar to investment decisions; their goal is not just to maximize an objective return. Instead, they think of the decision as relatively subjective, and therefore are often willing to compromise on effectiveness in order to satisfy their personal preferences. Studies 1 & 2 show that consumers are less likely to use effectiveness information for charitable decisions than for investment decisions. Study 3 shows that when searching for a charity, consumers are less likely to search by effectiveness ratings as compared to investments. Finally, Study 4 shows that presenting options in a relatively diverse (as opposed to uniform) choice set leads consumers to select a charity that fits their subjective preferences at the cost of maximizing effectiveness.

In Study 1 (N = 201), participants were told that they received a promotion for joining a local bank. In the Charity condition, the bank gave participants their choice to donate $100 to one of three fictitious charities, and a description of each charity was provided (e.g., “Main Street Housing Fund: This charity builds residential homes for those in need”). In the Investment condition participants could invest $100 into one of three fictitious companies (e.g., “Main Street Housing Corp: This company builds residential homes to sell to individuals”). Participants were also provided with effectiveness information for each option on a 0 to 100 scale, with higher numbers representing more effective options. For the charities, effectiveness was defined as doing “the most good for each dollar donated” whereas for the investments, effectiveness was defined as providing “the best return for each dollar invested”. Results show that in the Charity condition, significantly fewer participants chose the most effective option (45.5%) than in the Investment condition (64.0%; Z = 2.63, p = .009).

Study 2 (N = 214) replicates these results in a more controlled setting. First, we changed the scenario so that the decision was self-directed and was not tied to a financial institution. Specifically, participants were told to imagine that they sought either to donate or to invest $250 of their own money. Second, we presented participants with a wider range of options (15 options organized across 5 categories). Finally, to remove subtle differences that may exist between the investment and category options, participants only saw the name of option, its category, and effectiveness rating with no additional description (e.g., Arcadia Grocers, Category = Food, Rating = 85). Consistent with our previous findings, fewer participants in the Charity condition chose the most effective option (32.4%) than in the Investment condition (52.8%; Z = 3.02, p = .003).

We argue that for charitable decisions, participants weigh their subjective preferences greater than they do for investments. As a result, they are less likely to actively organize charity options by effectiveness ratings than they are for investment options. To test this, in Study 3, participants (N = 401) were given the same scenario as the previous study. However, before presenting participants with the choice set, we first asked whether they would like the options to be ordered by category or by effectiveness rating. Consistent with our expectation, fewer participants chose to order options by effectiveness in the Charity condition (67.8%) than in the Investment condition (83.4%; Z = -3.63, p < .001). Moreover, fewer participants in the Charity condition chose the most effective option (32.2%) than in the Investment condition (50.2%; Z = 3.68, p < .001). Finally, we find that the effect of domain condition (charity vs investment) on search mode decision (category vs rating) is mediated by participants’ feelings that their decision should reflect their personal beliefs (Indirect Effect = .68 95% CI = [.35, 1.08]), and is not mediated their desire for agency (Indirect Effect = .04 95% CI = [-.17, .25]), feelings of subjective knowledge in the domain (Indirect Effect = .04 95% CI = [-.01, .14]), or their belief that it is possible to rate options by effectiveness (Indirect Effect = -.03 95% CI = [-.13, .04]).

In Study 4 (N = 304), we manipulated whether participants were presented with a choice set in either a diverse or uniform set of options. We expect that when consumers compare charities from a diverse choice set, they choose less effective charities because the diverse set enables them to pick a cause that they personally prefer. Study 4 also utilizes a different effectiveness metric than the previous studies. Participants were presented with QALY (Quality Adjusted Life Years) ratings, which represents the number of fully healthy years of life that would be added to someone in need by a charity’s intervention.

Participants were randomly assigned to one of two conditions. In the Mixed Category condition, participants were presented with a list of three charities, each of which reflected a different category (Food, Medical, and International). In the Single Category condition, participants were presented with a list of three charities that all belonged to a single category (either Food, Medical, or International). We find that when choosing from a set of charities from different categories, participants picked the top rated option less often (69.3%)
than when choosing from a set of charities from a single category (84.1%; Z = 3.05, p = .002).

In sum, these studies show that when consumers decide where to donate their money, they often discount explicit effectiveness information, and instead consider their own subjective preferences.

REFERENCES